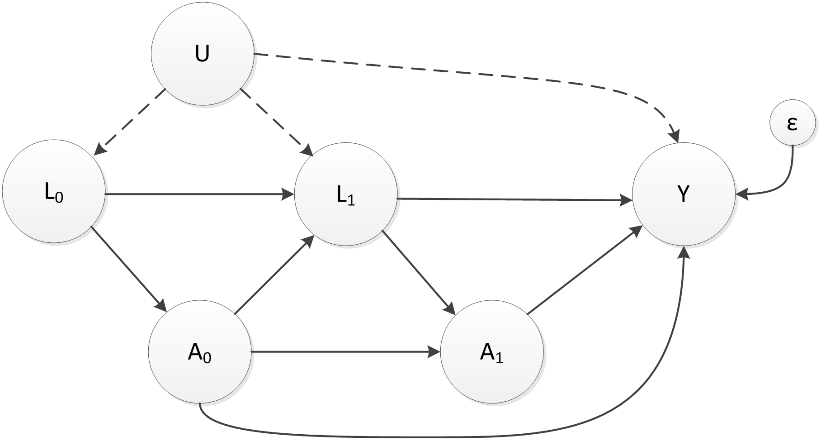
Using simulation to understand marginal structural modeling

What is the problem?



$$E(Y^1 - Y^0)$$

$$E(Y^1 - Y^0) = E(Y^1) - E(Y^0)$$

$$E(Y^0) \stackrel{?}{=} E(Y^0|A=0)$$

$$E(Y^1) \stackrel{?}{=} E(Y^1|A=1)$$

$$E(Y^0) = E(Y|A = 0)$$

$$E(Y^1) = E(Y|A = 1)$$

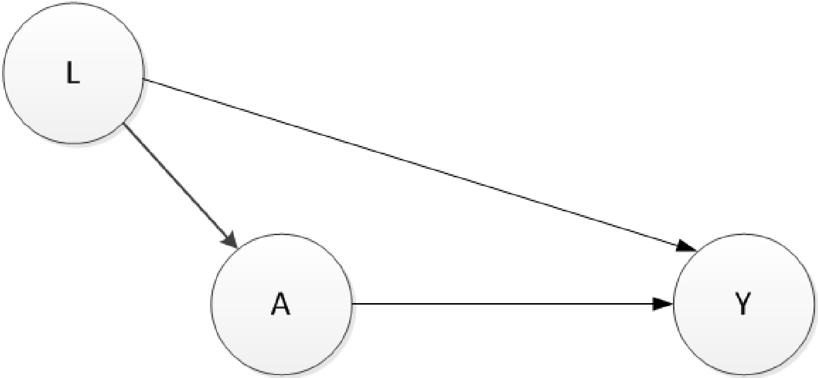
$$E(Y^0) \neq E(Y^0|A=0)$$

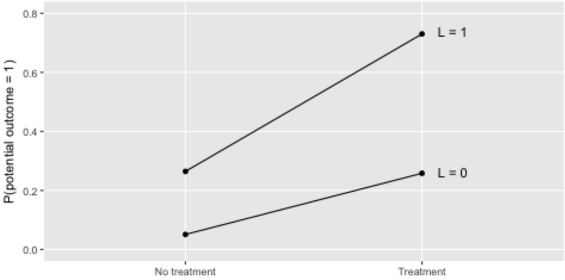
$$E(Y^1) \neq E(Y^1|A=1)$$

$$E(Y^0|L = l) = E(Y|A = 0 \text{ and } L = l)$$

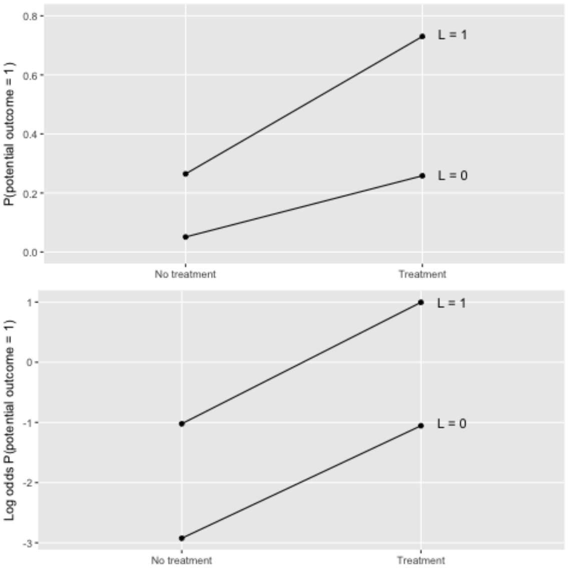
$$E(Y^1|L = l) = E(Y|A = 1 \text{ and } L = l)$$

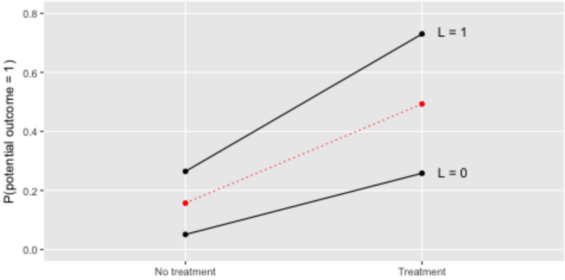
Simple simulation

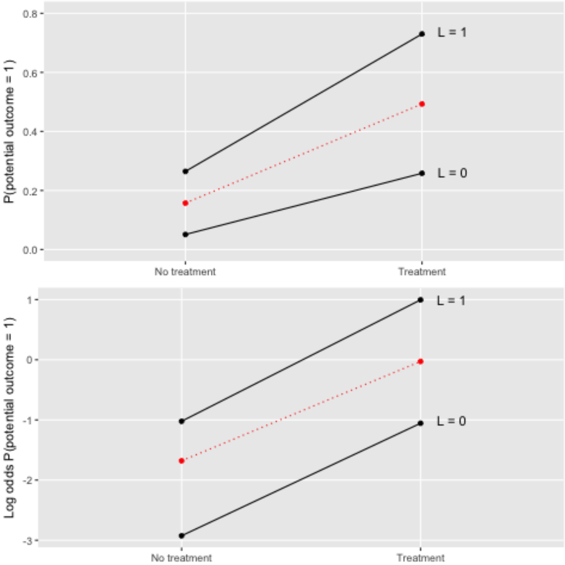




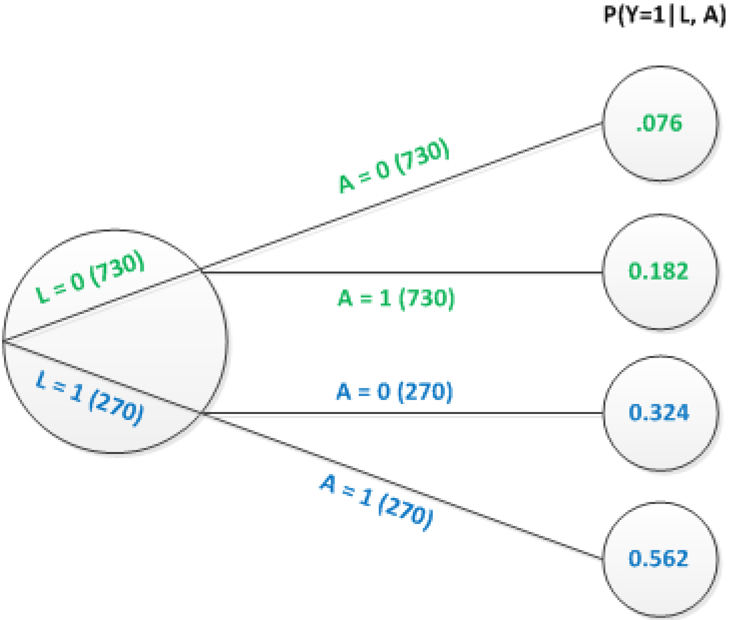
P(potential outcome = 1)
0.8
0.6
0.4
0.2
0.0
L = 1
L = 0
No treatment
Treatment


P(potential outcome = 1)
0.8
0.6
0.4
0.2
0.0
L = 1
L = 0
No treatment
Treatment
Log odds P(potential outcome = 1)
1
0
-1
-2
-3
L = 1
L = 0
No treatment
Treatment


P(potential outcome = 1)
0.8
0.6
0.4
0.2
0.0
L = 1
L = 0
No treatment
Treatment



P(potential outcome = 1)
0.8
0.6
0.4
0.2
0.0
L = 1
L = 0
No treatment
Treatment
Log odds P(potential outcome = 1)
1
0
-1
-2
-3
L = 1
L = 0
No treatment
Treatment


P(Y=1|L, A)
L = 0 (730)
L = 1 (270)
A = 0 (730)
A = 1 (730)
A = 0 (270)
A = 1 (270)
.076
0.182
0.324
0.562

Conditional log-odds ratios

$$LOR_{A=1 \textbf{ vs } A=0|L=0} = log\left(\frac{0.182/0.818}{0.076/0.924}\right) = log(2.705) = 0.995$$

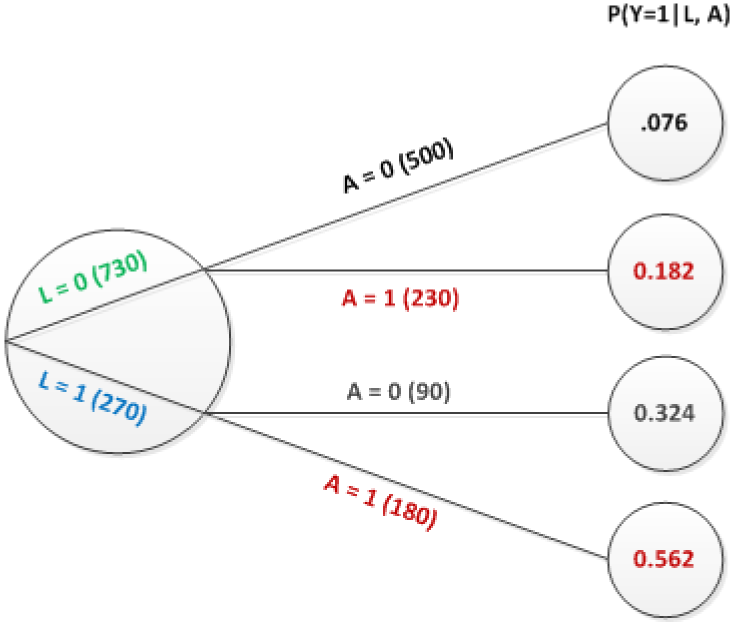
$$LOR_{A=1 \textbf{ vs } A=0|L=1} = log\left(\frac{0.562/0.438}{0.324/0.676}\right) = log(2.677) = 0.984$$

Marginal log-odds ratios

$$P(Y = 1|A = 0) = 0.73 \times 0.076 + 0.27 \times 0.324 = 0.143$$

$$P(Y = 1|A = 1) = 0.73 \times 0.182 + 0.27 \times 0.562 = 0.285$$

$$LOR_{A=1 \textbf{ vs } A=0} = log\left(\frac{0.285/0.715}{0.143/0.857}\right) = log(2.389) = 0.871$$


P(Y=1 | L, A)
L = 0 (730)
L = 1 (270)
A = 0 (500)
A = 1 (230)
A = 0 (90)
A = 1 (180)
.076
0.182
0.324
0.562

Crude log-odds ratio

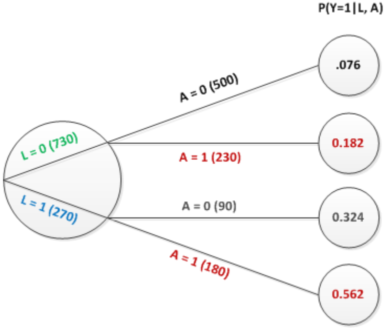
$$P(Y=1|A=0) = \frac{500 \times 0.076 + 90 \times 0.324}{500 + 90} = 0.114$$

$$P(Y=1|A=1) = \frac{230 \times 0.182 + 180 \times 0.562}{230 + 180} = 0.349$$

$$LOR_{A=1 \text{ vs } A=0} = log\left(\frac{0.349/0.651}{0.114/0.886}\right) = log(4.170) = 1.420$$

$$IPW = \frac{1}{P(A = a|L = l)}$$

What assumptions do we need to make?



Inverse probability weights

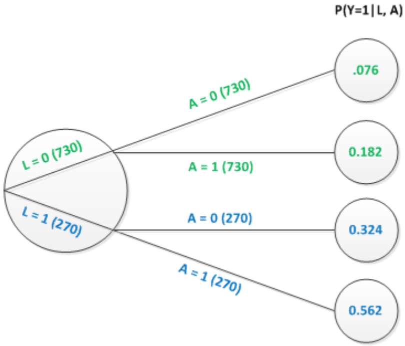
$$w_{00} = \frac{1}{P(A=0|L=0)} = \frac{1}{500/730} = \frac{730}{500} = 1.46$$

$$w_{01} = \frac{1}{P(A=1|L=0)} = \frac{1}{230/730} = \frac{730}{230} = 3.17$$

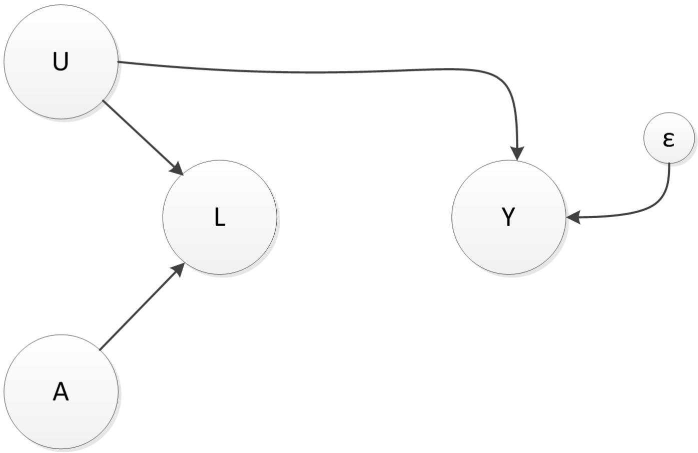
$$w_{10} = \frac{1}{P(A=0|L=1)} = \frac{1}{90/270} = \frac{270}{90} = 3.00$$

$$w_{11} = \frac{1}{P(A=1|L=1)} = \frac{1}{180/270} = \frac{270}{180} = 1.50$$

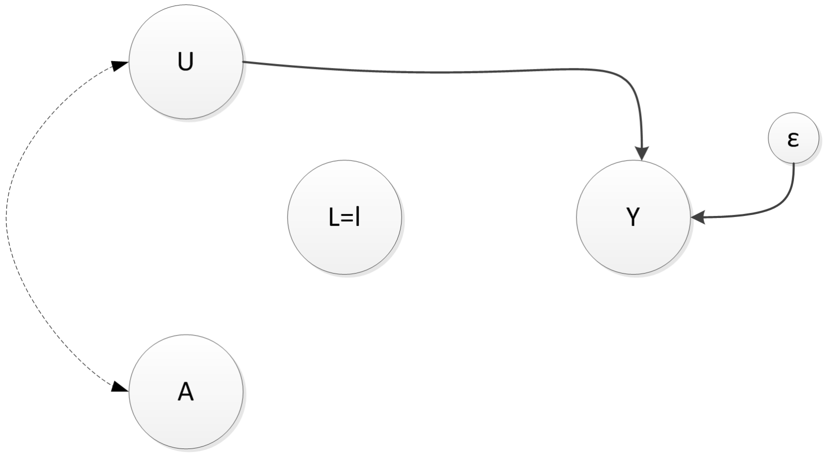
After applying weights



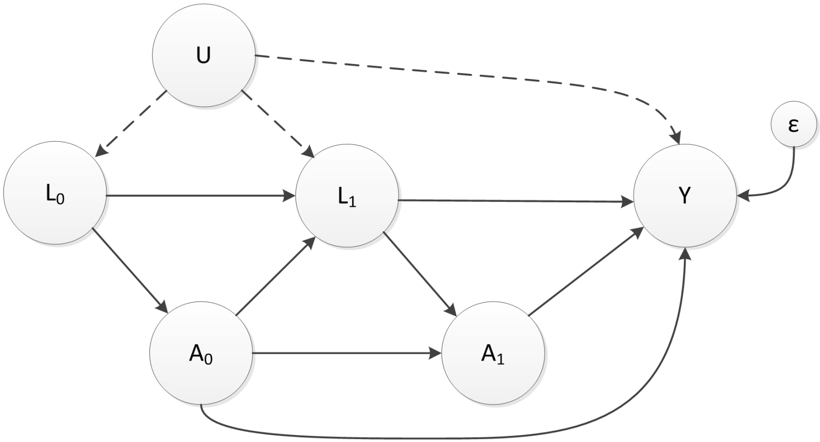
Simulation

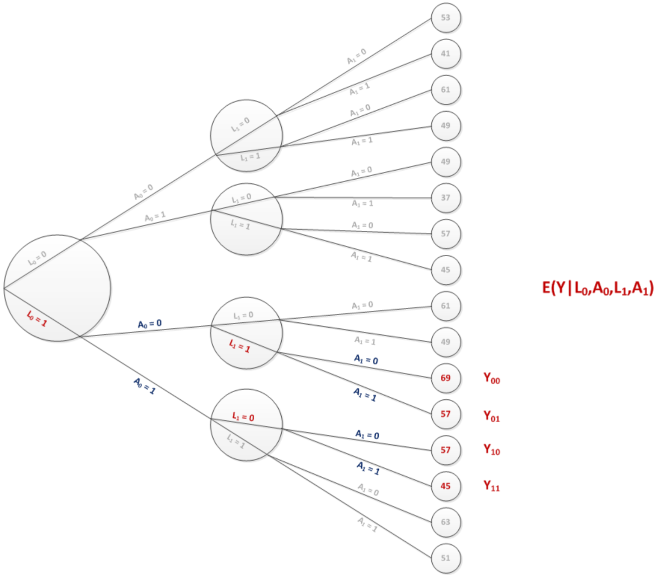


U
L
Y
ε
A


U
ε
L=l
Y
A

Simulation


U
ε
L_0
L_1
Y
A_0
A_1



L0 = 0
L0 = 1
A0 = 0
A0 = 1
A0 = 0
A0 = 1
L1 = 0
L1 = 1
A1 = 0
A1 = 1
53
41
61
49
49
37
57
45
61
49
69
57
57
45
63
51
Y00
Y01
Y10
Y11
E(Y|L0,A0,L1,A1)

Potential potential outcomes

$$E_1^i = Y_{10}^i - Y_{00}^i$$

$$E_2^i = Y_{01}^i - Y_{00}^i$$

$$E_3^i = Y_{11}^i - Y_{00}^i$$

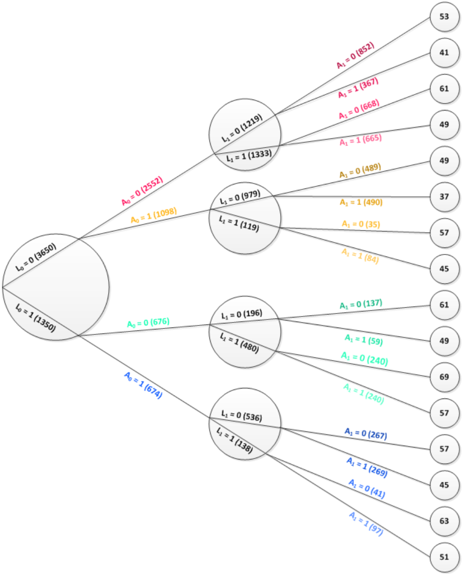
Starting, never stopping

$$E_1 = Y_{1111} - Y_{0000}$$

$$E_2 = Y_{0111} - Y_{0000}$$

$$E_3 = Y_{0011} - Y_{0000}$$

$$E_4 = Y_{0001} - Y_{0000}$$

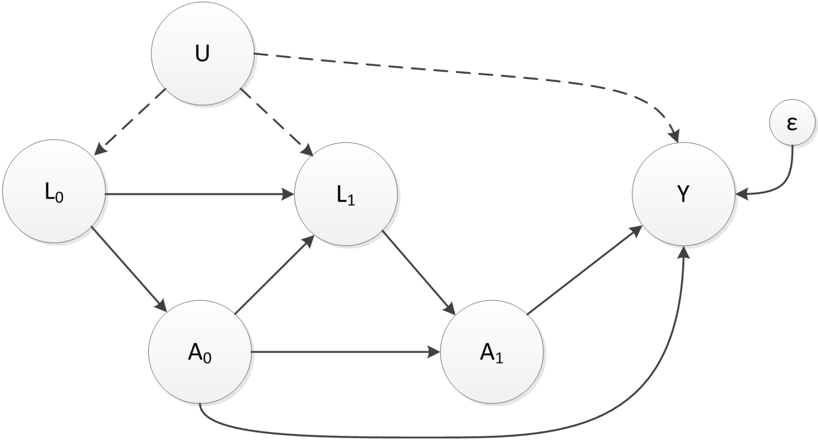

L0 = 0 (3650)
L0 = 1 (1350)
A0 = 0 (2552)
A0 = 1 (1098)
A0 = 0 (676)
A0 = 1 (674)
L1 = 0 (1219)
L1 = 1 (1333)
L1 = 0 (979)
L1 = 1 (119)
L1 = 0 (196)
L1 = 1 (480)
L1 = 0 (536)
L1 = 1 (138)
A1 = 0 (852)
A1 = 1 (367)
A1 = 0 (668)
A1 = 1 (665)
A1 = 0 (489)
A1 = 1 (490)
A1 = 0 (35)
A1 = 1 (84)
A1 = 0 (137)
A1 = 1 (59)
A1 = 0 (240)
A1 = 1 (240)
A1 = 0 (267)
A1 = 1 (269)
A1 = 0 (41)
A1 = 1 (97)
53
41
61
49
49
37
57
45
61
49
69
57
57
45
63
51

$E(Y|L_0,A_0,L_1,A_1)$

$$IPW = \frac{1}{P(A_0 = a_0, A_1 = a_1 | L_0 = l_0, L_1 = l_1)}$$

$$IPW = \frac{1}{P(A_0 = a_0, A_1 = a_1 | L_0 = l_0, L_1 = l_1)}$$

$$= \frac{1}{P(A_0 = a_0 | L_0 = l_0) \times P(A_1 = a_1 | L_0 = l_0, A_0 = a_0, L_1 = l_1)}$$


U
ε
L_0
L_1
Y
A_0
A_1

Simulation

Key references

Havercroft, W. G., and V. Didelez. “Simulating from marginal structural models with time-dependent confounding.” Statistics in medicine 31.30 (2012): 4190-4206.

Robins, James M., Miguel Angel Hernan, and Babette Brumback. “Marginal structural models and causal inference in epidemiology.” (2000): 550-560.